Journal of Dentistry 135 (2023) 104585

Contents lists available at ScienceDirect

Journal of Dentistry

journal homepage: www.elsevier.com/locate/jdent

Gaze patterns of dentists while evaluating bitewing radiographs☆

Lubaina T. Arsiwala-Scheppach [a,b,*], Nora Castner [c], Csaba Rohrer [a], Sarah Mertens [a], Enkelejda Kasneci [d], Jose Eduardo Cejudo Grano de Oro [a], Joachim Krois [a,b], Falk Schwendicke [a,b]

[a] Department of Oral Diagnostics, Digital Health and Health Services Research, Charité - Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Germany

[b] ITU/WHO Focus Group AI on Health, Topic Group Dental Diagnostics and Digital Dentistry, Switzerland

[c] Department of Computer Science, University of Tuebingen, Tuebingen, Germany

[d] Department of Computer Science, Technical University of Munich, Germany

A R T I C L E I N F O

Keywords:
Dental radiography
Dentistry
Eye tracking
Medical image inspection
Scanpath
Visual search

A B S T R A C T

Objectives: Understanding dentists' gaze patterns on radiographs may allow to unravel sources of their limited accuracy and develop strategies to mitigate them. We conducted an eye tracking experiment to characterize dentists' scanpaths and thus their gaze patterns when assessing bitewing radiographs to detect primary proximal carious lesions.

Methods: 22 dentists assessed a median of nine bitewing images each, resulting in 170 datasets after excluding data with poor quality of gaze recording. Fixation was defined as an area of attentional focus related to visual stimuli. We calculated time to first fixation, fixation count, average fixation duration, and fixation frequency. Analyses were performed for the entire image and stratified by (1) presence of carious lesions and/or restorations and (2) lesion depth (E1/2: outer/inner enamel; D1–3: outer-inner third of dentin). We also examined the transitional nature of the dentists' gaze.

Results: Dentists had more fixations on teeth with lesions and/or restorations (median=138 [interquartile range=87, 204]) than teeth without them (32 [15, 66]), $p<0.001$. Notably, teeth with lesions had longer fixation durations (407 milliseconds [242, 591]) than those with restorations (289 milliseconds [216, 337]), $p<0.001$. Time to first fixation was longer for teeth with E1 lesions (17,128 milliseconds [8813, 21,540]) than lesions of other depths ($p = 0.049$). The highest number of fixations were on teeth with D2 lesions (43 [20, 51]) and lowest on teeth with E1 lesions (5 [1, 37]), $p<0.001$. Generally, a systematic tooth-by-tooth gaze pattern was observed.

Conclusions: As hypothesized, while visually inspecting bitewing radiographic images, dentists employed a heightened focus on certain image features/areas, relevant to the assigned task. Also, they generally examined the entire image in a systematic tooth-by-tooth pattern.

1. Introduction

Bitewing radiographs are a standard complementary method for detecting and staging carious lesions, showing higher sensitivity and similar specificity than visual-tactile examination, especially for proximal lesions [1]. Notably, dentists often do not achieve the theoretically possible high sensitivity, as a recent large-scale review and meta-analysis demonstrated, with more than every second lesion remaining non-detected in mean [2]. This oversight may be routed in the way dental experts assess imagery, such as (bitewing) radiographs.

Efficient and thorough inspection of medical images leads to faster feature recognition and better clinical reasoning [3–5] which is crucial for medical professionals, such as radiologists and dentists, who regularly interpret a high volume of these images. When assessing these images, professionals employ both a heightened focus to certain features and prior knowledge, leading to a context dependent gaze known as *scanpath* which comprises of 'fixations' (attentional information) and 'saccades' (transitions to attentional areas) [6,7]. Much of the previous

☆ Acknowledgements: Not applicable

* Corresponding author at: Department of Oral Diagnostics, Digital Health and Health Services Research, Charité - Universitätsmedizin Berlin, Aßmannshauser Str. 4-6, 14197 Berlin, Germany.

E-mail address: lubaina.arsiwala@charite.de (L.T. Arsiwala-Scheppach).

https://doi.org/10.1016/j.jdent.2023.104585
Received 7 February 2023; Received in revised form 15 May 2023; Accepted 7 June 2023
Available online 8 June 2023

literature on gaze patterns has focused on the comparison between experts and novices, confirming that experts are usually faster, but nevertheless more accurate than novices when assessing imagery [8,9], which is realized by a number of aspects. For instance, experts show a shorter time to first fixation on relevant areas (e.g., an anomaly) than novices, suggesting that their experience provides sophisticated shortcuts [5,10–16]. Experts also have more fixations per image, and fixations of longer duration on relevant areas compared to irrelevant areas, which can be attributed to reducing extraneous attentional processing [8,17–20]. Moreover, image content has been shown to have a significant impact on expert eye movements [5,8,21,22]. Obvious anomalies do not require as many fixations for experts as inconspicuous ones [5,23,24]. In mammograms, dental computed tomography (CT) scans, and dental panoramic radiographs, for example, experts have fewer fixations for more obvious anomalies than novices but overall, more fixations, particularly on more subtle anomalies [22,25,26]. The nature of the first fixation and area revisits is further affected by the image content; on dental periapical radiographs, the presence of restorations has been found to affect both factors [27]. Last, image type and purpose affect gaze patterns. For certain types of medical images, like chest CT and dental panoramic radiographs, experts form a global representation of the content at a glance [28,29], and then usually employ a systematic scanning pattern over small image areas. For example, gaze patterns with an inward spiral focus on the peripheral areas first and the dental areas second [12,13] or circular scan patterns have been identified [30–32]. In contrast, in brain CT scans, gaze patterns are saliency-driven, which creates a focal-then-global search strategy [33]. On dental periapical radiographs, tooth-by-tooth viewing is common [27]. A more detailed overview about medical gaze patterns can be found elsewhere [3,34,35], while generally, gaze patterns for dental image interpretation is not well understood.

The aim of the present study was to characterize and describe dentists' gaze patterns when identifying primary proximal carious lesions on bitewing radiographs of the permanent dentition. Our hypothesis was that dentists would demonstrate different gaze patterns for different types of image content.

2. Materials and methods

2.1. Study design

This evaluation is nested within a randomized, controlled, non-blinded, clustered cross-over, superiority trial with an allocation ratio of 1:1 [36], assessing the impact of an artificial intelligence (AI) software on detection of carious lesions. The trial was not conducted during clinical care and on actual patients, but on retrospectively sampled imagery material obtained from patients treated at Charité – Universitätsmedizin Berlin and collected between the years 2016 and 2018, which was randomly assessed with and without assistance from the AI software. The trial was registered at Deutsches Register Klinischer Studien (DRKS00022357). Ethical approval was provided by the Charité – Universitätsmedizin Berlin (EA/144/20). During the study, we recorded dentists' gaze data, and here we present the gaze patterns of the control group (i.e., dentists not using AI). Written informed consent was obtained from all participating dentists.

2.2. Participants and image data

Recruitment of participants and study conduct took place between October 2020 and January 2021. Participants were dentists employed at the dental hospital of Charité – Universitätsmedizin Berlin or in private practices in Berlin, Germany. The study was performed either in the dental hospital of Charité – Universitätsmedizin Berlin or at the private practice of the participants. All participating dentists had more than two years of clinical experience (i.e., had finished postgraduate education according to German insurance law). Exclusion criteria for the participants were not being clinically active any longer or having no regular experience with caries detection (e.g., orthodontists or oral surgeons). A total of 22 dentists were recruited.

Bitewing radiographs of primary teeth or those where assessment was deemed impossible were excluded. This resulted in one hundred and forty bitewing radiographs of the permanent dentition, with at least the crowns of one dental arch being detectable, being included. Most of the images (63%) were generated using radiographic machines from the manufacturer Dentsply Sirona (Bensheim, Germany), mainly Orthophos XG; the rest using Dürr Dental machines (Bietigheim-Bissingen, Germany).

The establishment of the reference test used for confirming the presence of carious lesions and denoting their depth is described elsewhere [37]. Briefly, four experts independently labeled proximal lesions in a pixel-wise manner on all images (i.e., each expert labeled 140 bitewing images) using an in-house custom-built annotation tool in dimly lit rooms on diagnostic screens and under standardized conditions. All labels on all images were reviewed and curated (additions, deletions, corrections) by a fifth expert dentist, who had the chance of contacting the experts for further discussion. No formal metrics of agreement (e.g., inter- and intra-examiner agreement) were collected and no triangulation with any clinical records was performed. The reference test was constructed by the union of all annotated areas for each carious lesion on every image. Note that this labeling process has been employed previously and is one of several options for establishing a reference test on fuzzy data (i.e. where a "hard" reference test like histopathology is unavailable). The lesion depth was defined by two independent examiners in agreement, as follows; E1 denoted lesions into the outer half of the enamel, E2 those into the inner enamel half but not extending into the dentin, D1 those not extending deeper than the outer 1/3rd of the dentin, D2 those not extending deeper than the outer 2/3rd of the dentin, and D3 those extending beyond the outer 2/3rd of the dentin.

2.3. Eye tracker

To record gaze data, the remote eye tracker *SmartEye Aurora* running at 60 Hz was positioned under a monitor which had full high-definition resolution of 1920 * 1080 pixels. The study room was dimly lit and the participants were unconstrained and positioned approximately 70 cm from the tracker. For the participants from private clinics, the study investigator brought the monitor to their clinic and the experiment was carried out in a dimly lit room in the clinic. An initial 9-point calibration and validation were performed. Recalibration was done if the software indicated that the calibration quality was poor. Gaze data was collected for the whole duration of the study and then pre-processed using the *iMotions* software (version 8.2.22899.4). Event detection was performed using the *iMotions* implementation of the I-VT algorithm, with a minimum fixation duration of 60 milliseconds (ms) and a velocity threshold of 30°/second. The current analysis used the fixations reported from the software, which are interpolated between the left and the right eye. We interpret fixations as the areas of attentional focus related to the stimuli presented on the screen.

2.4. Gaze data

Each of the 22 dentists received a set of ten randomly selected bitewing radiographs, data from which has been examined in the present study. Details on the randomization of the radiographs from the pool of 140 bitewing radiographs have been described in our previous study [36]. Data collection resulted in 221 datasets from the participants viewing bitewing radiographs. As three participants unintentionally examined one image twice, we excluded the first time they viewed the image, as it was too short for proper investigation. To ensure gaze data quality, we removed datasets with an average reported gaze signal quality lower than 0.60 (on a scale of 0.0 being the lowest and 1.0 being the highest quality) [38]; 43 datasets were excluded by this criterion. Stimulus presentation error resulted in the additional removal of five datasets. The exclusion of data was non-systematic i.e., no bias was introduced in the study results. Overall and finally, 170 datasets were included in the current analysis. Each dentist viewed a median of nine bitewing radiograph images and each image was viewed by a median of two dentists.

2.5. Outcomes and covariates

We analyzed fixation behavior of dentists while visually inspecting bitewing radiographic images. We begin by defining each outcome and its unit of measurement. The time to first fixation indicates the amount of time that it takes for a dentist to fixate on a specific area of interest (AOI) from the onset of stimulus and is measured in ms [39]. Fixation count provides information on how many times a dentist returned their gaze to a particular AOI and is measured as a numeric count [39]. Average fixation duration quantifies how long on average a fixation lasted for and is measured in ms [39].

General fixation metrics like average fixation duration (ms) offer insights into how professionals holistically process bitewing radiographs. Additionally, the fixation count related to the time per specific task indicates how quickly the relevant information is extracted. Since viewing times per image were variable, an additional measure of fixation frequency per second i.e., number of fixations per second, was calculated. Time to first fixation (ms) as well as fixation count and average fixation duration (ms) in relevant regions are indicators of efficient information retrieval. The relevant regions for this study were the teeth, proximal carious lesions, and restorations visible on the bitewing images. For marking the teeth, an (unpublished) in-house tooth detection model, whose findings had been validated by an experienced dentist for each bitewing image, was employed. The carious lesions and restorations were established by the reference test devised for this study, as laid out earlier.

To further investigate the procedural aspect of dentists' gaze, we looked at the transitions of the scanpaths, i.e., how often the gaze transitioned to a neighboring tooth (e.g., from tooth 24 to tooth 25) or somewhere else instead (e.g., from tooth 26 to tooth 37). To account for image dependent patterns, transition matrices were created for ten images that were viewed by at least three dentists.

2.6. Statistical analysis

Participant characteristics such as age and gender were recorded and used for descriptive analyses. All analyses were performed for the total dataset (i.e., overall) and stratified by presence of carious lesions and/or restorations, and carious lesion depth (E1/2: outer/inner enamel; D1–3: outer to inner third of dentin). The relevant variables exhibited non-normal distributions and thus were summarized using median and inter-quartile range (IQR) and were analyzed using non-parametric tests. Differences in each gaze metric between relevant groups were tested using the Wilcoxon rank sum test and Kruskal-Wallis test, as appropriate, where level of significance was set to $p<0.05$. Missing data was not imputed.

To account for any possible spatial offsets in the gaze data, AOIs were given an extra pixel padding based on their relative pixel area. The teeth were large enough to simply have the bounding boxes. Based on our study setup, the average size of a bounding box for a tooth was 325.98 by 234.97 pixels, which is approximately 9.2 cm by 6.6 cm on the monitor relative to the participant. The bounding boxes of restorations were given a pixel padding of 3° of visual angle (which approximates to 129 pixels for our specific setup), and bounding boxes of carious lesions were given 3-, 7-, or 10° padding based on whether their area was on the larger, medium, or smaller side of the lesion area distribution, respectively. For fixation behavior analysis, we counted fixations that land in overlapping AOIs as a *hit* in both AOIs. All statistical analyses and data management were performed using Python (version 3.8 and above) and R (version 4.0.3, www.r-project.org).

3. Results

3.1. Gaze patterns

Six female and 16 male dentists participated; their mean age was 38 years (range: 27 years to 60 years). The dentists spent a median of 49 seconds (IQR = 34, 72) per bitewing image, with a median fixation count of 167 per bitewing (IQR = 127, 212) and a median fixation frequency of 3 fixations per second (IQR = 3, 4). Upon evaluating the fixation frequency in context with the average fixation duration, longer fixation durations corresponded to slower fixation frequencies, whereas shorter durations corresponded to faster fixation frequencies, see Fig. 1. Meanwhile, the distribution of the fixation frequency in relation to the average fixation dispersion (mean = 0.36°; standard deviation = 0.04°) showed that for slower fixation frequencies, lesser image area was inspected, and more image area was covered at faster frequencies, see Fig. 2.

3.2. Gaze patterns related to carious lesions and restorations

When looking at dentists' viewing patterns on teeth with and without carious lesions and/or restorations, there were 364 teeth with lesions and 481 teeth with restorations (overall 581 teeth, as some showed both), and 365 teeth without. Regarding gaze on the teeth, in 79% (129/170) of datasets a tooth with lesions and/or restorations was looked at first. Median time to first fixation showed no significant difference for teeth with lesions and/or restorations (median = 359 ms, IQR = 181, 674) and teeth without lesions and/or restorations (median = 384 ms, IQR = 236, 612), $p = 0.68$.

Focusing on the teeth with lesions and/or restorations, we noted that a majority of dentists initially fixated on teeth with restorations (72% (125/170) datasets) and not lesions. Median time to first fixation was also shorter for teeth with restorations (median = 1275 ms, IQR = 501, 4075) than for teeth with lesions (median = 6598 ms, IQR = 2945, 20,669), $p<0.001$. Dentists had significantly more fixations on teeth with lesions and/or restorations (median = 138, IQR = 87, 204) than teeth without lesions and/or restorations (median = 32, IQR = 15, 66), $p<0.001$. There was a large difference in the number of fixations on teeth with restorations (median = 47, IQR = 19, 100) compared with those with lesions (median = 17, IQR = 6, 32), $p<0.001$. Average fixation duration on teeth with (median = 337 ms, IQR = 251, 413) and without lesions and/or restorations (median = 308 ms, IQR = 227, 367) did not differ. Notably, teeth with lesions had longer fixation durations (median = 407 ms, IQR = 242, 591) compared with those with restorations (median = 289 ms, IQR = 216, 337), $p<0.001$.

Regarding gaze patterns on teeth with carious lesions of different depths (Table 1), the longest time to first fixation was for teeth with E1 lesions (median = 17,128 ms, IQR =8813, 21,540) as compared to teeth with lesions of other depths ($p = 0.049$) and they were also looked at first in only 7% of cases. Teeth with E2 lesions were looked at first the most i.e., in 40% of cases. Regarding fixation counts on teeth with lesions ($n = 364$), the highest number of fixations were on teeth with D2 lesions (median = 43, IQR = 20, 51) and lowest on teeth with E1 lesions (median = 5, IQR = 1, 37), $p<0.001$. Average fixation duration showed less variability with the longest fixation duration being for teeth with E1 lesions (median = 530 ms, IQR = 468, 664) and the shortest for teeth with D3 lesions (median = 310 ms, IQR = 255, 417), $p = 0.49$.

3.3. Scanpath and transitional behavior

We also investigated the nature of transition of the scanpaths, i.e., how often the gaze transitioned to a neighboring tooth versus non-neighboring tooth. We noted that the highest number of transitions

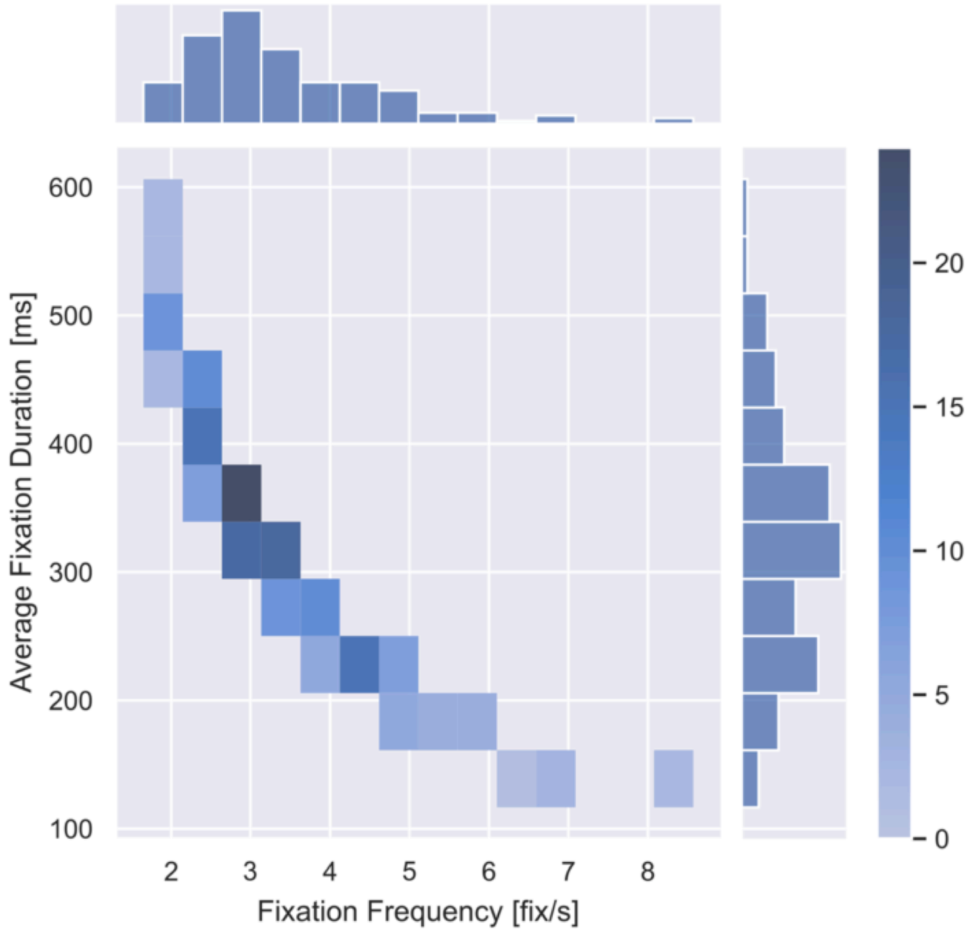

Fig. 1. Distribution of fixation frequency (number of fixations per second) in relation to average fixation duration (milliseconds) while evaluating the bitewing radiographic images.

were to the neighboring tooth (n = 14,119 transitions), and the second most frequent transitions were to the second next tooth (n = 1105 transitions); Table 2. Often, there were transitions to a tooth in the opposite jaw (n = 2635 transitions); however, to a lesser extent.

In order to confirm this lateral tooth-by-tooth visual inspection, transition matrices were created for exemplary ten images (images A to J) that were viewed by at least three dentists to control for image dependent scanpath patterns. Fig. 3 depicts transition matrices from three of these images (images A to C) and the remaining seven images (images D to J) are shown in Supplementary Fig. 1. In Fig. 3, the transition matrices showed that dentists generally examined the images tooth by tooth (lighter colors along the diagonal in the image indicated higher number of transitions to the neighboring tooth). More interestingly, there was no preference for left to right inspection or right to left inspection, which would be illustrated by one side of the diagonal being lighter than the other. Note that the dentists also looked outside of the bitewing image when using the digital viewing platform and its functionalities.

We also qualitatively examined the individual scanpaths for the subset of images based on their semantic information in Fig. 3, which we have labeled as A, B, and C. Image A had no carious lesions, but restorations (including root-canal fillings) on teeth 25, 35, 36, and 37. The scanpaths for image A offer an almost ideal visualization of systematic lateral tooth-by-tooth visual inspection, and the transition matrix for image A quantitatively highlights that the neighboring teeth with restorations had the highest transitions (transitions from tooth 36 to 35, from 36 to 37, and from 37 to 36). Image B had lesions of depth E2 on teeth 16 and 17 and restorations on teeth 46 and 47. The scanpaths for image B show similar systematic strategies, whereas the transition matrix points out that the highest transitions were related to teeth with lesions (between teeth 16 and 17) and restorations (between teeth 46 and 47). Image C had lesions and restorations only in the maxilla; E2 lesions on teeth 24 and 26, an E1 lesion on tooth 25, D1 on tooth 27, and a restoration on tooth 26. The scanpaths for image C qualitatively appear to be the least systematic in comparison to images A and B, with overall higher concentrations of fixations and longer saccades (the lines connecting fixations), and slightly more transitions between the jaws. The transition matrix nevertheless confirms that the majority of transitions were to neighboring teeth. The highest number of transitions were from tooth 37 to 36, neither of which have any lesions or restorations. The majority of these transitions were from the dentist denoted as ‘Dentist 3’ (blue in Fig. 3).

4. Discussion

Gaze patterns for dental radiographs are not well understood, which is why the present study aimed to characterize dentists’ gaze patterns when evaluating bitewing radiographs for primary proximal carious lesions. As hypothesized, dentists demonstrated different gaze patterns for different types of image content, in context of the task they were given. Our results are suggestive of a systematic search strategy being employed by most dentists with respect to the nature of the task assigned to them, i.e., identifying proximal caries in bitewing radiographs. To begin with, longer fixation duration was related to slower fixation frequency which in turn was related to lesser image area being inspected. These patterns are indicative of how dentists interpret the image

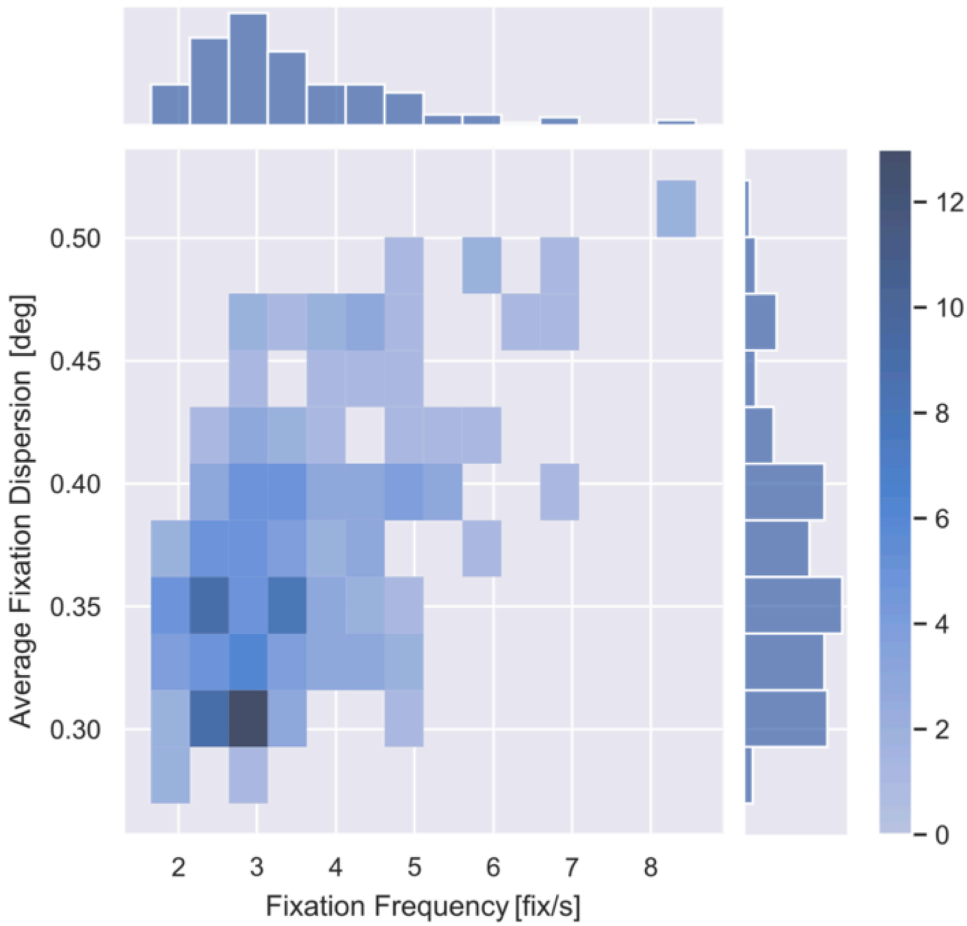

Fig. 2. Distribution of fixation frequency (number of fixations per second) in relation to average fixation dispersion (degrees) while evaluating the bitewing radiographic images.

Table 1
Distribution of gaze characteristics on teeth with carious lesions, stratified by depth of lesion.

Gaze metrics	Carious lesion depth					p-value
	E1	E2	D1	D2	D3	
Time to First Fixation, median (IQR), milliseconds	17,128 (8813, 21,540)	9398 (3850, 33,388)	8390 (2955, 20,420)	5146 (3021, 6987)	3300 (2567, 9082)	0.049
Total Fixation Count, median (IQR)	5 (1, 37)	10 (2, 22)	15 (10, 27)	43 (20, 51)	25 (18, 31)	<0.001
Average Fixation Duration, median (IQR), milliseconds	530 (468, 664)	381 (227, 614)	447 (199, 604)	486 (285, 569)	310 (255, 417)	0.49

The lesion depth was defined as follows; E1 denoted lesions into the outer half of the enamel, E2 those into the inner enamel half but not extending into the dentin, D1 those not extending deeper than the outer 1/3rd of the dentin, D2 those not extending deeper than the outer 2/3rd of the dentin, and D3 those extending beyond the outer 2/3rd of the dentin.
The p-values apply to the entire table row.
IQR, inter-quartile range.

Table 2
Number of transitions of the dentists' gaze across teeth when viewing a bitewing radiographic image.

From the current tooth to …	Number of transitions of the dentists' gaze
The next tooth in the same jaw	14,119
2 teeth away in the same jaw	1105
3 teeth away in the same jaw	207
4 teeth away in the same jaw	22
Any tooth in the opposite jaw	2635

information during visual inspection, where more rapid fixation over larger areas can be interpreted as more global or ambient scanning pattern and slower, more thorough attention to smaller areas as more focal scanning pattern. In our study the observed gaze patterns allude to visual inspection strategies that are predominantly shorter dispersions, with more concentrated durations (higher) and frequencies (slower). Another aspect highlights the relevance of the task designated to the dentists: most dentists first looked at teeth with caries and/or restorations and fixated more on them compared to teeth without. Also, fixation durations were longer on teeth with lesions than restorations and though this aligns with the nature of the task, it was also indicative of thorough inspection of the teeth that they determined to bear carious lesions. Additionally, the transitional nature of the scanpath patterns (see

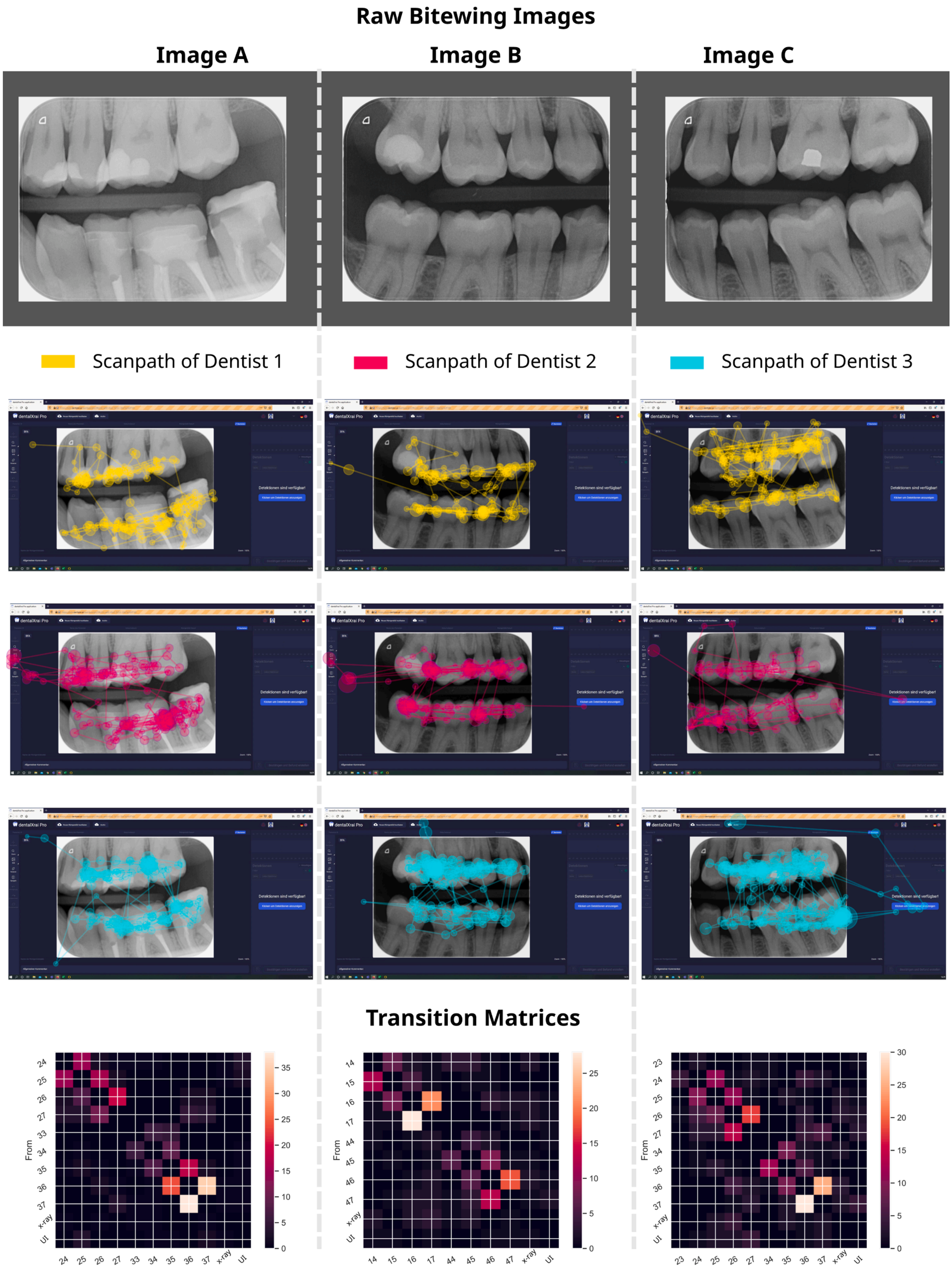

Fig. 3. Original images, their scanpath data, and transition matrices of three bitewing radiographic images which generally elicited systematic search patterns in the study. UI, user interface.

Table 2), i.e., most frequent transitions being to the neighboring teeth, suggested a systematic lateral tooth-by-tooth visual inspection of teeth in one jaw, before moving onto the other jaw, in comparison to other approaches such a top-to-bottom visual inspection or unsystematic assessment. This pattern was similarly linked to the nature of the task (i.e., identifying proximal caries) and image type (i.e., bitewing radiograph) [40]. More fixations were noted on teeth with restorations than lesions. Although this was unexpected, there were possibly more fixations to thoroughly inspect and determine that there are in fact no carious lesions on teeth with existing restorations (e.g., secondary carious lesions even though the detection of these lesions was not the task given to the dentists). The result of these thorough evaluations was supported by the previously reported high accuracy of 0.93 (95% confidence interval: 0.92, 0.95) of dentists for this task in this study [36]. When stratified by carious lesion depth, our results are in contrast with those from other studies, especially regarding teeth with E1 lesions. Other studies have reported that obvious and easy-to-spot anomalies do not require as many fixations for experts than inconspicuous and harder-to-detect anomalies [5,23,24]. In our results, teeth with E1 lesions, which are incipient and thus difficult to identify, had the lowest count of average fixations and longest time to first fixation. One potential reason for this may lie with the dentists' performance in diagnosing the lesions. A prior evaluation of the dentists' performance in this study showed a low sensitivity of 0.64 (95% confidence interval: 0.53, 0.74) in diagnosing early enamel lesions, suggesting that the dentists were likely to miss these incipient lesions, thereby leading to the lower average count of fixations [36].

The applications of gaze pattern analysis in dentistry are multifold since the field heavily relies on imagery for clinical and research tasks. In education, it offers the potential to assess a student's learning progress in real-time and thus provides opportunities to adapt the stimulus material based on current aptitude [40]. On the other hand, it can also be used to assess the effectiveness of training modules. Additionally, examination of the scanpath visualizations along with the transition matrices can offer interesting insights into the respective cognitive strategies of a specific dentist. Future research should further investigate the role of image complexity on gaze patterns. Additionally, gleaning deeper insights into how professionals extract data from medical images may serve in building better AI-supported diagnostic tools and take us a step further in the direction of explainable AI.

This study has a number of strengths and limitations. First, it uses a range of outcomes to comprehensively characterize the gaze patterns of dentists. We will engage in understanding the impact of AI on gaze patterns and the diagnostic process in another evaluation of the yielded trial data, as it seems relevant to unravel how exactly AI changes dentists' diagnostic behavior. Second, and as limitation, our cohort of bitewing radiographic images and dentists was limited and selected. As described, the imagery stemmed from two machines and one clinical center and thus the generalizability of the results on other imagery cannot be expected. Also, a small and selected sample of dentists was enrolled; the sample was younger than the average German dentist, mainly situated in an urban clinic or practice environment. Third, the identification of whether a tooth had a carious lesion and/or restoration was built on a reference test constituted by five experts; no further validation (e.g., histology) was performed. It can be expected that even five experts and their verdict may not always yield "the ground truth", a caveat we accepted. Similarly, the lesion depth was determined by two examiners jointly, which may come with limited robustness as well.

5. Conclusions

Using gaze pattern analysis, we demonstrated that when dentists visually inspected dental bitewing radiographic images their gaze was specific for different types of image content and was determined partially by the context of the task they were assigned. Dentists predominantly employed shorter dispersions and more focus on areas of the image relevant to the task of identifying proximal carious lesions. Additionally, a systematic tooth-by-tooth gaze was commonly observed in this study. The assessment of one's gaze pattern is an efficient and non-invasive method to collect objective data on the complex interplay of one's cognition and education/training for accomplishing a given task. Further research in this direction can help us glean insights into the causes for dentists' limited accuracy (for example in diagnosing proximal caries on bitewing radiographs) and thus develop strategies to improve their clinical performance.

Clinical significance

Analysis of dentists' gaze patterns offers objective insights into how they extract data from different types of dental images. Such characterization of clinicians' gaze patterns may determine their cognitive and training status, which are integral to achieving successful clinical outcomes, and thus identify ways to optimize their data extraction methods.

Supplementary Fig. 1. Original images, their scanpath data, and transition matrices of the seven bitewing radiographic images used in the study. UI, user interface.

CRediT authorship contribution statement

Lubaina T. Arsiwala-Scheppach: Software, Validation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Nora Castner:** Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Writing – review & editing, Visualization. **Csaba Rohrer:** Methodology, Software, Formal analysis, Data curation. **Sarah Mertens:** Investigation, Writing – review & editing. **Enkelejda Kasneci:** Conceptualization, Methodology, Resources. **Jose Eduardo Cejudo Grano de Oro:** Methodology, Software, Formal analysis. **Joachim Krois:** Conceptualization, Project administration. **Falk Schwendicke:** Conceptualization, Methodology, Resources, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jdent.2023.104585.

References

[1] F. Schwendicke, M. Tzschoppe, S. Paris, Radiographic caries detection: a systematic review and meta-analysis, J. Dent. 43 (8) (2015) 924–933.

[2] T. Walsh, R. Macey, P. Riley, A.M. Glenny, F. Schwendicke, H.V. Worthington, J. E. Clarkson, D. Ricketts, T.L. Su, A. Sengupta, Imaging modalities to inform the detection and diagnosis of early caries, Cochrane Database Syst. Rev. 3 (2021), CD014545.

[3] A. Ganesan, M. Alakhras, P.C. Brennan, C. Mello-Thoms, A review of factors influencing radiologists' visual search behaviour, J. Med. Imaging Radiat. Oncol. 62 (6) (2018) 747–757.

[4] K.A. Ericsson, A.C. Lehmann, Expert and exceptional performance: evidence of maximal adaptation to task constraints, Annu. Rev. Psychol. 47 (1996) 273–305.

[5] A. van der Gijp, C.J. Ravesloot, H. Jarodzka, M.F. van der Schaaf, I.C. van der Schaaf, J.P.J. van Schaik, T.J. Ten Cate, How visual search relates to visual diagnostic performance: a narrative systematic review of eye-tracking research in radiology, Adv. Health Sci. Educ. Theory Pract. 22 (3) (2017) 765–787.

[6] D. Noton, L. Stark, Scanpaths in saccadic eye movements while viewing and recognizing patterns, Vision Res. 11 (9) (1971) 929–942.

[7] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, J. van de Weijer. Eye tracking: A comprehensive guide to methods and measures, 1st, Oxford University Press, 2011.

[8] A. Gegenfurtner, E. Lehtinen, R. Säljö, Expertise differences in the comprehension of visualizations: a meta-analysis of eye-tracking research in professional domains, Educ. Psychol. Rev. 23 (4) (2011) 523–552.
[9] R. Nakashima, C. Watanabe, E. Maeda, T. Yoshikawa, I. Matsuda, S. Miki, K. Yokosawa, The effect of expert knowledge on medical search: medical experts have specialized abilities for detecting serious lesions, Psychol. Res. 79 (5) (2015) 729–738.
[10] C.F. Nodine, H.L. Kundel, S.C. Lauver, L.C. Toto, Nature of expertise in searching mammograms for breast masses, Acad. Radiol. 3 (12) (1996) 1000–1006.
[11] S. Brams, G. Ziv, O. Levin, J. Spitz, J. Wagemans, A.M. Williams, W.F. Helsen, The relationship between gaze behavior, expertise, and performance: a systematic review, Psychol. Bull. 145 (10) (2019) 980–1027.
[12] H.L. Kundel, C.F. Nodine, E.F. Conant, S.P. Weinstein, Holistic component of image perception in mammogram interpretation: gaze-tracking study, Radiology 242 (2) (2007) 396–402.
[13] E.A. Krupinski, Visual scanning patterns of radiologists searching mammograms, Acad. Radiol. 3 (2) (1996) 137–144.
[14] C. Lindsey, G. Alastair, S. Janak, G. Swamy, S. Hazel, T. Andoni, The assessment of stroke multidimensional CT and MR imaging using eye movement analysis: does modality preference enhance observer performance? Proc. SPIE 7627 (2010), 76270B.
[15] S. Mallett, P. Phillips, T.R. Fanshawe, E. Helbren, D. Boone, A. Gale, S.A. Taylor, D. Manning, D.G. Altman, S. Halligan, Tracking eye gaze during interpretation of endoluminal three-dimensional CT colonography: visual perception of experienced and inexperienced readers, Radiology 273 (3) (2014) 783–792.
[16] A.L. Warren, T.L. Donnon, C.R. Wagg, H. Priest, N.J. Fernandez, Quantifying novice and expert differences in visual diagnostic reasoning in veterinary pathology using eye-tracking technology, J. Vet. Med. Educ. 45(3) 295–306.
[17] J.J. Topczewski, A.M. Topczewski, H. Tang, L.K. Kendhammer, N.J. Pienta, NMR spectra through the eyes of a student: eye tracking applied to NMR Items, J. Chem. Educ. 94 (1) (2017) 29–37.
[18] H. Haider, P.A. Frensch, Eye movement during skill acquisition: more evidence for the information-reduction hypothesis, J. Exp. Psychol. Learn. Mem. Cogn. 25 (1999) 172–190.
[19] M.A. Just, P.A. Carpenter, A capacity theory of comprehension: individual differences in working memory, Psychol. Rev. 99 (1) (1992) 122–149.
[20] S.P. Marshall, Identifying cognitive state from eye metrics, Aviat. Space Environ. Med. 78 (5 Suppl) (2007) B165–B175.
[21] E.M. Kok, A.B.H. De Bruin, S.G.F. Robben, J.J.G. Van Merriënboer, Looking in the same manner but seeing it differently: bottom-up and expertise effects in radiology, Appl. Cogn. Psychol. 26 (2012) 854–862.
[22] D.P. Turgeon, E.W. Lam, Influence of experience and training on dental students' examination performance regarding panoramic images, J. Dent. Educ. 80 (2) (2016) 156–164.
[23] G. Wood, K.M. Knapp, B. Rock, C. Cousens, C. Roobottom, M.R. Wilson, Visual expertise in detecting and diagnosing skeletal fractures, Skeletal Radiol. 42 (2) (2013) 165–172.
[24] T. Donovan, D. Litchfield, Looking for cancer: expertise related differences in searching and decision making, Appl. Cogn. Psychol. 27 (1) (2013) 43–49.
[25] K. Suwa, A. Furukawa, T. Matsumoto, T. Yosue, Analyzing the eye movement of dentists during their reading of CT images, Odontology 89 (1) (2001) 54–61.
[26] F. Alamudun, H.J. Yoon, K.B. Hudson, G. Morin-Ducote, T. Hammond, G. D. Tourassi, Fractal analysis of visual search activity for mass detection during mammographic screening, Med. Phys. 44 (3) (2017) 832–846.
[27] B.P. Hermanson, G.C. Burgdorf, J.F. Hatton, D.M. Speegle, K.F. Woodmansey, Visual fixation and scan patterns of dentists viewing dental periapical radiographs: an eye tracking pilot study, J. Endod. 44 (5) (2018) 722–727.
[28] C.F. Nodine, C.R. Mello-Thoms, The nature of expertise in radiology. SPIE, 2000, pp. 859–895.
[29] N. Castner, J. Frankemölle, C. Keutel, F. Huettig, E. Kasneci, LSTMs can distinguish dental expert saccade behavior with high "plaque-urracy", in: 2022 Symposium On Eye Tracking Research and Applications, Association for Computing Machinery, Seattle, WA, USA, 2022, p. 8. Article.
[30] T. Grünheid, D.A. Hollevoet, J.R. Miller, B.E. Larson, Visual scan behavior of new and experienced clinicians assessing panoramic radiographs, J. World Fed. Orthod. 2 (1) (2013) e3–e7.
[31] T. Drew, M.L. Vo, A. Olwal, F. Jacobson, S.E. Seltzer, J.M. Wolfe, Scanners and drillers: characterizing expert visual search through volumetric images, J. Vis. 13 (10) (2013).
[32] E. Mercan, L.G. Shapiro, T.T. Brunye, D.L. Weaver, J.G. Elmore, Characterizing diagnostic search patterns in digital breast pathology: scanners and drillers, J. Digit. Imaging 31 (1) (2018) 32–41.
[33] H. Matsumoto, Y. Terao, A. Yugeta, H. Fukuda, M. Emoto, T. Furubayashi, T. Okano, R. Hanajima, Y. Ugawa, Where do neurologists look when viewing brain CT images? An eye-tracking study involving stroke cases, PLOS One 6 (12) (2011) e28928.
[34] Z. Gandomkar, C. Mello-Thoms, Visual search in breast imaging, Br. J. Radiol. 92 (1102) (2019), 20190057.
[35] S.E. Fox, B.E. Faulkner-Jones, Eye-tracking in the study of visual expertise: methodology and approaches in medicine, Frontline Learn. Res. 5 (3) (2017) 43–54.
[36] S. Mertens, J. Krois, A.G. Cantu, L.T. Arsiwala, F. Schwendicke, Artificial intelligence for caries detection: randomized trial, J. Dent. 115 (2021), 103849.
[37] A.G. Cantu, S. Gehrung, J. Krois, A. Chaurasia, J.G. Rossi, R. Gaudin, K. Elhennawy, F. Schwendicke, Detecting caries lesions of different radiographic extension on bitewings using deep learning, J. Dent. 100 (2020), 103425.
[38] K. Holmqvist, S.L. Orbom, I.T.C. Hooge, D.C. Niehorster, R.G. Alexander, R. Andersson, J.S. Benjamins, P. Blignaut, A.M. Brouwer, L.L. Chuang, K. A. Dalrymple, D. Drieghe, M.J. Dunn, U. Ettinger, S. Fiedler, T. Foulsham, J.N. van der Geest, D.W. Hansen, S.B. Hutton, E. Kasneci, A. Kingstone, P.C. Knox, E.M. Kok, H. Lee, J.Y. Lee, J.M. Leppanen, S. Macknik, P. Majaranta, S. Martinez-Conde, A. Nuthmann, M. Nystrom, J.L. Orquin, J. Otero-Millan, S.Y. Park, S. Popelka, F. Proudlock, F. Renkewitz, A. Roorda, M. Schulte-Mecklenbeck, B. Sharif, F. Shic, M. Shovman, M.G. Thomas, W. Venrooij, R. Zemblys, R.S. Hessels, Eye tracking: empirical foundations for a minimal reporting guideline, Behav. Res. Methods (2022).
[39] B. Farnsworth, 10 Most used eye tracking metrics and terms, 2022. https://imotions.com/blog/learning/10-terms-metrics-eye-tracking/. (Accessed 07/05/2023).
[40] N. Castner, T.C. Kuebler, K. Scheiter, J. Richter, T. Eder, F. Huettig, C. Keutel, E. Kasneci, Deep semantic gaze embedding and scanpath comparison for expertise classification during OPT viewing, ACM Symposium on Eye Tracking Research and Applications, Association for Computing Machinery, Stuttgart, Germany (2020) 18.